Math 3339

Section 27204
MWF 10-11:00am AAAud 2

Bekki George
bekki@math.uh.edu
639 PGH

Office Hours:
M & Th noon – 1:00 pm & T 1:00 – 2:00 pm
and by appointment

Let X and Y be two normal random variables with $X \sim Norm(\mu_X, \sigma_X)$ and $Y \sim Norm(\mu_Y, \sigma_Y)$. A bivariate normal distribution depends on five parameters, μ_x, μ_y, $\sigma_x > 0$, $\sigma_y > 0$, and $\rho \in (-1,1)$. $\boxed{\rho \text{ is the correlation.}}$ ρ

Theorem 3 *Let X_i, $i = 1, \cdots, n$ be independent normally distributed random variables $X_i \sim Norm(\mu_i, \sigma_i)$. Then $X_1 + X_2 + \cdots + X_n$ is normally distributed with mean $\mu_1 + \mu_2 + \cdots + \mu_n$ and variance $\sigma_1^2 + \cdots + \sigma_n^2$.*

$\sigma = \sqrt{\sigma_1^2 + \sigma_2^2 + \cdots + \sigma_n^2}$

Not independent

$Y = 5 + 3X_1 + 4X_2$

Theorem 4 *Let X_1 and X_2 have a bivariate normal distribution with parameters μ_1, μ_2, σ_1, σ_2 and ρ. Let a, b_1, and b_2 be constants and let $Y = a + b_1 X_1 + b_2 X_2$. Then Y is normally distributed with mean $a + b_1\mu_1 + b_2\mu_2$ and variance $b_1^2\sigma_1^2 + b_2^2\sigma_2^2 + \boxed{2b_1b_2\rho\sigma_1\sigma_2}$.*

$Y = 2X + 1$ $\quad E[Y] = 2E[X] + 1$

Hw 7 problems 6.5 #1 & 2:

1. John's projected annual income after graduation is normally distributed with mean $80,000 and standard deviation $10,000. Sally's is normally distributed with mean $85,000 and standard deviation $12,000. If their incomes are independent, what is the probability that their combined income exceeds $180,000?

$Y = X_1 + X_2 \quad P(Y > 180000)$

2. What is the standard deviation of their combined income if their incomes have a bivariate normal distribution and the correlation between them is 0.4? What is it if the correlation is -0.4?

$\rho = .4$

What is the mean and standard deviation for #1?

$\mu_Y = 80000 + 85000 = 165{,}000$

$\sigma_Y = \sqrt{\sigma_Y^2} = \sqrt{10{,}000^2 + 12000^2}$

How about #2?

$\sigma_Y = \sqrt{10{,}000^2 + 12000^2 + 2(.4)(10000)(12000)}$

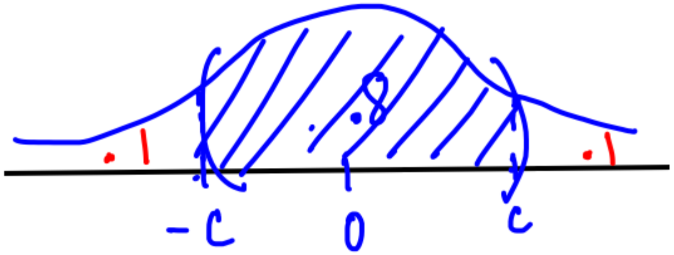
$\mu_Y = 165000$

Basic Properties of Confidence Intervals

standard normal

c = Z-score
how many st. dev. from mean

Recall: Find a value of c so that $P(-c < Z < c) = 0.8$



qnorm(.1) = -c

qnorm(.1 + .8) = c

c = 1.282

$\underbrace{}_{1 - .8 = .2}$

In general, how would you find a value of c so that $\underbrace{P(-c < Z < c)}_{P(|Z| < c)} = 1 - \alpha$?

"within c"

qnorm(α/2) = -c

qnorm(α/2 + 1 - α) = qnorm(1 - α/2) = c

mean, variance, proportion

The confidence interval consists of two parts: an interval and a confidence level.
The **interval** is our $\boxed{\text{estimate}}$ ± margin of error
The **confidence level** gives the probability that the method produces an interval that covers the parameter.

A 95% confidence level says, "We got these numbers by a method that gives correct results 95% of the time."

When working with means for normal distributions, the margin of error is the range of values above and below the sample statistic. It shows how accurate we believe our guess is, based on the variability of the estimate.

$$m = z^*\left(\frac{\sigma}{\sqrt{n}}\right)$$

standard error

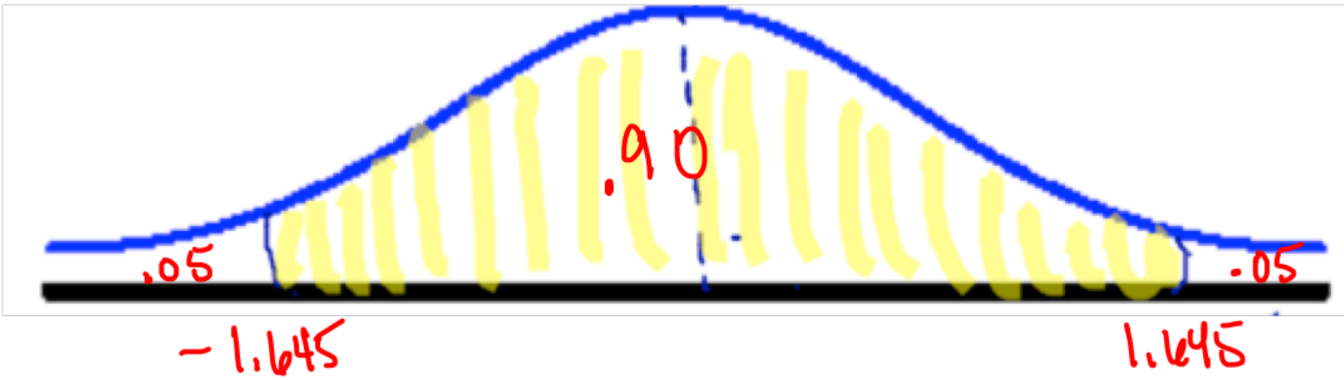
when working w/ distr. of $\bar{x}$ recall $\sigma_{\bar{x}} = \sigma/\sqrt{n}$

c or z-score associated w/ level of confidence

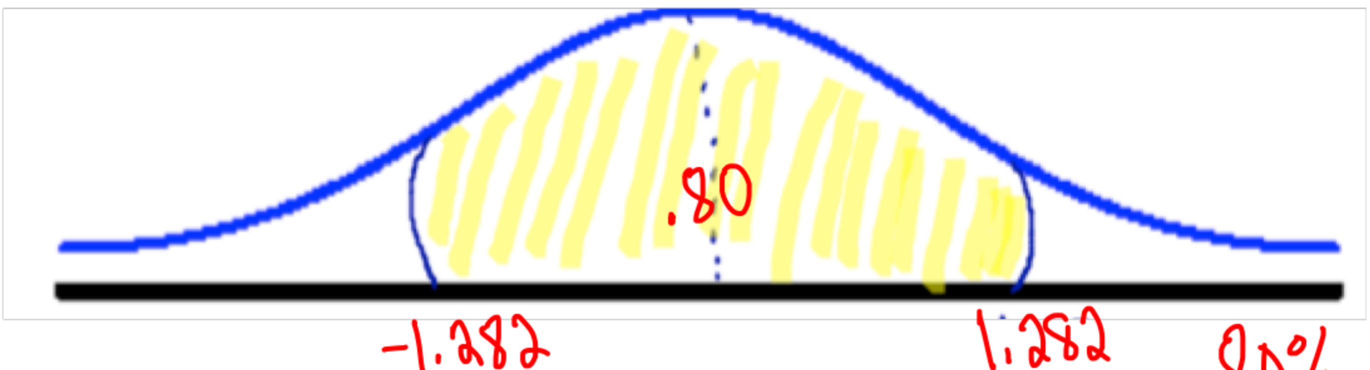
In general, for any distribution, the margin of error = Critical value x Standard error of the statistic

The figure below shows the general situation for any confidence level C and illustrates how z^* is calculated:

90% conf. interval



qnorm(.05)
= -1.645
qnorm(.95)
= 1.645



80% conf. interval

c for which $P(-c < z < c) = 1 - \alpha$

A $(1-\alpha)$ Confidence interval for μ:

90% ⇒ each tail is 5% ($\alpha/2$)

$$\left[\bar{x} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right]$$

represents the interval for which $1-\alpha$ is the probability that μ is in the interval.

Assumptions:

1. Population distribution is normal
2. σ is known

Warnings about using the formula $\boxed{\bar{x} \pm z^*\left(\frac{\sigma}{\sqrt{n}}\right)}$:

1. The data must be a SRS from the population.
2. Do not use the formula for probability sampling designs more complex than a SRS.
3. There is no correct method for inference of badly produced data.
4. Outliers can have a large effect on the confidence level. Correct or remove outliers before computing the interval.
5. The confidence interval will be slightly off when the sample size is small and the population is not normal. As long as $n \geq 15$, the confidence level is not greatly disturbed by nonnormal populations unless extreme outliers or very strong skewness are present.

30 is better

6. You must know the standard deviation of the population.

(20, 35)

Definition of Confidence Interval

A **level C confidence interval** for a parameter is an interval computed from sample data by a method that has probability C of producing an interval containing the true value of the parameter.

When asked for an interpretation of a confidence interval, your response should follow this outline:

"We are ______% confident that the true mean lies in the interval 20 to 35."

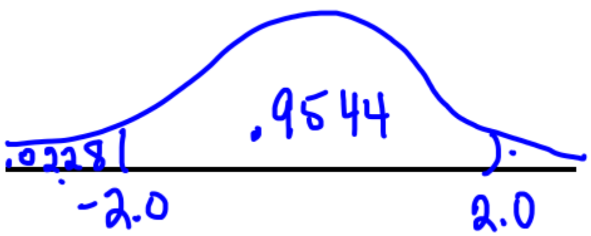
When asked for an interpretation of a confidence **level**, your response should follow this outline:

90%

"If we were to construct many 100 intervals with this method, 90 % of them would contain the true mean."

$N(\mu, \sigma^2)$ $\sigma = 2$

Example: A random sample of size $n = 25$ from the distribution $N(\mu, 4)$ results in a sample mean $\bar{x} = 13.2$. Give a 95.44% confidence interval for the mean μ.



$1 - .9544 = .0456$

$.0456/2 = .0228$

$z^* = 2.0$

-qnorm(.0228)

qnorm(.9772)

$\bar{x} \pm z^* \left(\sigma/\sqrt{n}\right)$

$13.2 \pm 2.0 \left(2/\sqrt{25}\right)$

$13.2 \pm .8$

margin of error

true mean = population mean

μ

(12.4, 14.0)

We are 95.44% confident the true mean is between 12.4 and 14.0

Larger sample size $\Rightarrow$ margin of error smaller

Determination of Sample Size

Suppose that we want to choose n large enough so that we are 95% confident that our population mean μ is within 4 units of the sample mean $\bar{x}$. How would we do this?

$z^* = 1.96$

$\alpha = .05$

$-qnorm(.05/2)$

$\text{M.E.} \le 4$

$\bar{x} \pm 4$

$z^* \cdot \frac{\sigma}{\sqrt{n}}$

$$1.96\left(\frac{\sigma}{\sqrt{n}}\right) \le 4$$

$$\frac{1.96\sigma}{4} \le \sqrt{n}$$

$$\left(\frac{1.96\sigma}{4}\right)^2 \le n$$

ME

+ integer

ex: $12.1 \le n$

$n = 13$

In general, if we want to be $(1-\alpha)100$ percent confident that μ is within h of $\bar{x}$, we create a confidence interval

ME

$$\left[\bar{X} - z_{\alpha/2}\frac{\sigma}{\sqrt{n}}, \bar{X} + z_{\alpha/2}\frac{\sigma}{\sqrt{n}}\right], \text{ where } \boxed{m = z_{\alpha/2}\frac{\sigma}{\sqrt{n}}} \leq h$$

and we solve for n.

Note: The *width* of the confidence interval above is $2m$.

Example: Suppose that the daily yield for a paint manufacturing process is normally distributed with standard deviation 3 tons. How many sample observations do we have to obtain if we want to be 95% certain that our estimate is within 1 ton of the actual mean yield μ?

$\sigma = 3$ $\quad n = ?$ $\quad 1-\alpha = .95$ $\quad m \leq 1$

$\alpha = .05$

$z^* = 1.96$

$1.96\left(\frac{3}{\sqrt{n}}\right) \leq 1$

$5.88 \leq \sqrt{n}$

$34.57 \leq n$

$n = 35$

Popper 17

1. What effect does increasing the sample size have on the width of a confidence interval? Does it make it longer, shorter, or stay the same?
   a. Longer
   b. Shorter
   c. Stays the same

A simple random sample of 25 new food cans indicated that the mean diameter was 3.13 inches. It is known that the standard deviation of all cans is .04 inches. We want to construct a 95% confidence interval based on this data.

2. What is z*?
   a. 1.96 ←
   b. 1.64
   c. 2.33
   d. 2.05
   e. none of these

$$\begin{cases} n = 25 \\ \bar{x} = 3.13 \\ \sigma = .04 \end{cases}$$

$\bar{x} \pm$ m.E.

3. What is the margin of error?
    a. 0.0784
    b. 0.0031
    c. 0.0157
    d. 0.0131
    e. none of these

4. Give the 95% confidence interval
    a. (3.1143, 3.1457)
    b. (3.0332, 3.2338)
    c. (3.1015, 3.1519)
    d. none of these

5. A